Problem 1.2

Here are the first lines of a professor's data set at the end of a statistics course:

```
NAME            MAJOR   POINTS  GRADE

ADVANI, SURA    COMM    397     B
BARTON, DAVID   HIST    323     C
BROWN, ANNETTE  BIOL    446     A
CHIU, SUN       PSYC    405     B
CORTEZ, MARIA   PSYC    461     A
```

What are the individuals and the variables in these data? Which variables are categorical and which are quantitative?

Problem 1.12

The number of deaths amoung persons aged 15 to 24 years in the United States in 1999 due to the eight leading causes of death for this age group were: accidents, 13,602; homicide, 4989; suicide 3885; cancer, 1724; heart disease, 1048; congential defects, 430; respiratory disease, 208; AIDS, 197.

    a. Make a bar graph of this data

    b. What additional information do you need to make a pie chart?

Problem 1.16

Make another stemplot of the percent of residents aged 25 to 34 in each of the states by splitting the stems in the plot from the previous exercise (see exercise 1.15 below). Which plot do you prefer? Why?

Problem 1.24

There is some evidence that increasing the amount of calcium in the diet can lower blood pressure. In a mdeical experiment one group of men was

given a daily calcium supplement, while a control group received a placebo (a dummy pill). The seated systolic blood pressure of all the mean was measured before the treatments began and again after 12 weeks. The blood pressure dstributions in the two groups should have been similar at the beginning of the experiment. Here are the intial blood pressure readings for the two groups:

```
Calcium Group
107  110  123  129  112  111  107  112  136  102


Placebo Group
123  109  112  102   98  114  119  112  110  117  130
```

Make a back-to-back stemplot of these data. Does your plot show any major differences in the two groups before the treatments began? In particular, are the centers of the two blood pressure distributions close together?

Problem 1.30

The following table gives the survival time in days of 72 guinea pigs after they were injected with tubercle bacilli in a medical experiment. Make a suitable graph and describe the shape, center and spread of the distribution of survival times. Are there any outliers

```
TABLE: Survival times (days) of guinea pigs in a medical experiment

43   45   53   56   56   57   58   66   67   73
74   79   80   80   81   81   81   82   83   83
84   88   89   91   91   92   92   97   99   99
100  100  101  102  102  102  103  104  107  108
109  113  114  118  121  123  126  128  137  138
139  144  145  147  156  162  174  178  179  184
191  198  211  214  243  249  329  380  403  511
522  598
```

Problem 1.46

Stemplots help you find the five-number summary because they arrange the observations in increasing odere. Exercise 1.15 (see below) inncludes a stemplot of the percent of residents aged 25 to 34 in each of the 50 states;

    a. Find the five-number summary of this distribution.

    b. does teh $1.5 \times IQR$ criterion flag Montana and Wyoming as suspected outliers?

    c. How much does the median change if you omit Montana and Wyoming?

Exercise 1.15 (for reference only, do not attempt)

Here is a stemplot of the percents of residents aged 25 to 34 in each of the 50 states. The stems are whole percents and the leaves are tenths of a percent.

```
10 | 9
11 | 0
12 | 1 3 4 4 6 7 7 8 8 9
13 | 0 0 1 2 4 5 5 5 6 6 7 8 9 9 9 9
14 | 1 1 2 2 2 3 4 4 4 4 5 7 8 9
15 | 2 4 4 7 8 9 9 9
```

    a. Montana and Wyoming have the smallest percents of young adults, perhaps because they lack job opportunities. What are the percetns for these two states?

    b. Ignoring Montana and Wyoming, describe the shate, center and spread of this distribution.